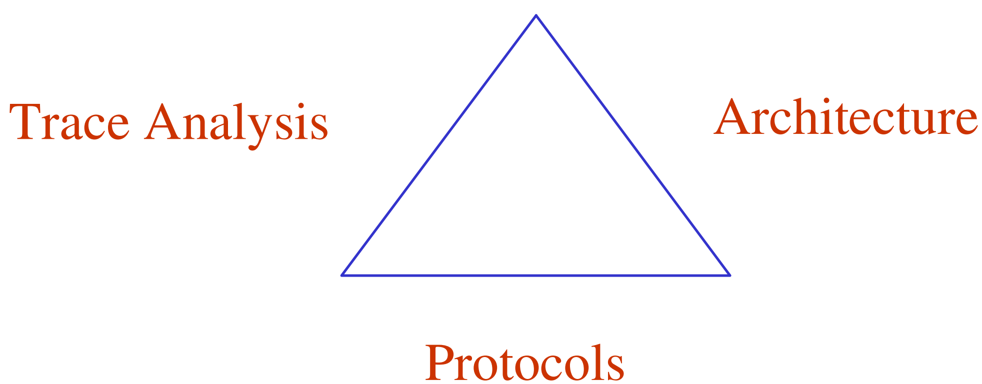
# Reducing Overhead in Flow-Switched Networks: An Empirical Study of Web Traffic


Anja Feldmann, Jennifer Rexford, Ramon Caceres

AT&T Labs-Research; Florham Park, NJ

# IP Traffic over Connections

## Problems in existing Internet

- Growing traffic load is straining router capacity
- Best-effort service provides little protection between users

## Solution with flow switching

- Efficient, high-speed switching in hardware
- Quality-of-service at the connection level

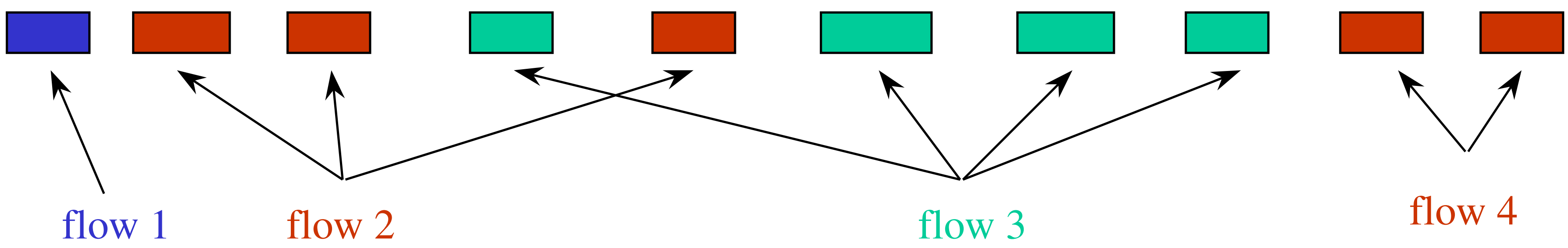
## Technical challenges

- Grouping related IP packets into flows
- Understanding the signatures of the dominant application Web
- Characterizing the connection overheads

## Performance results

- Detailed study of 1 week of trace data from T1 connecting AT&T Murray Hill to the Internet
- New methods to reduce connection overheads

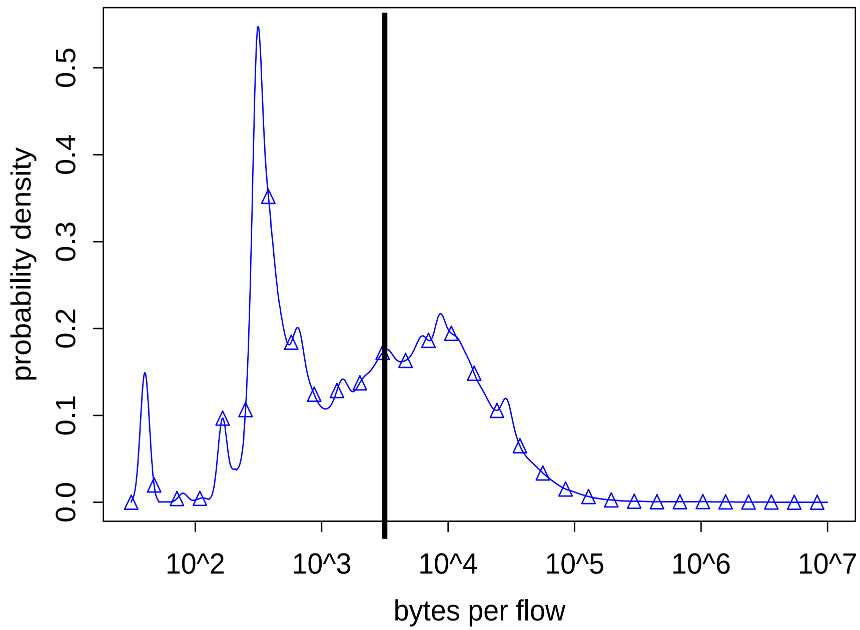
# Grouping IP Packets Into Flows




- Group packets with the "same" address
  - Application-level: single transfer from web server to client
  - Host-level: multiple transfers from server to client
  - Subnet-level: multiple transfers to a group of clients
- Group packets that are "close" in time
  - 60-second spacing between consecutive packets
- Client filter to "signal" flow definition and QoS

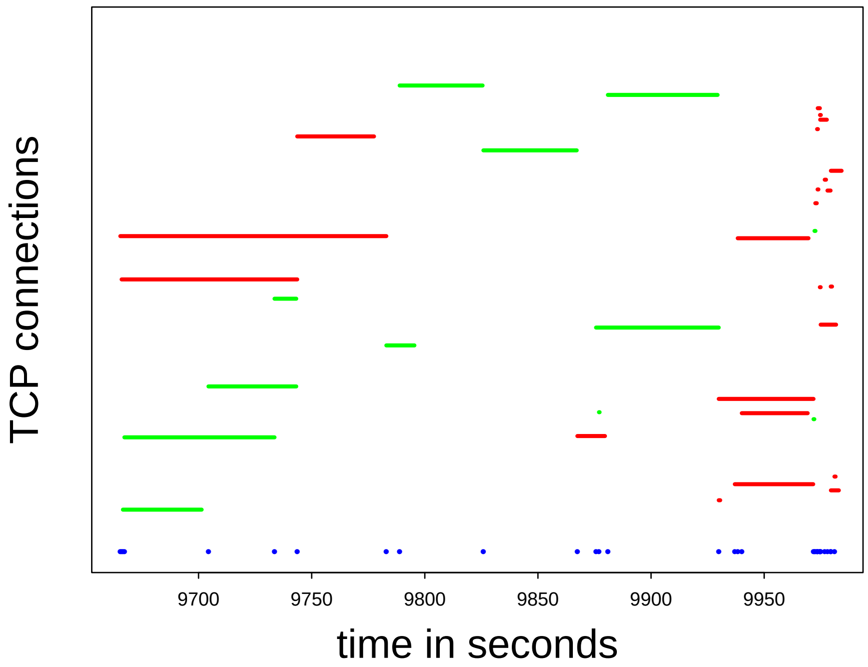
**Study flow size distributions (#bytes, #packets, #seconds)**

# Short-Lived vs. Long-Lived Flows Web Flows




- Many very short flows (30% are less than 300 bytes)
  - 18% “not modified”, 10% various error messages, e.g., “not found”
- Many medium-sized flows (short web transfers)
- Most bytes belong to long flows (large images, files)
- **Flow densities are signatures**

# IP-flows Within a Web Session




- Each line corresponds to a TCP connection (start time to end time)
- Same color implies same Web server

- parallel request to same Web server
- consecutive requests to same Web server
- bursts of requests
- highly variable durations

## Host-to-host flows

- on average 5.5 application flows per host-to-host flow
- shift of probability density to longer flows

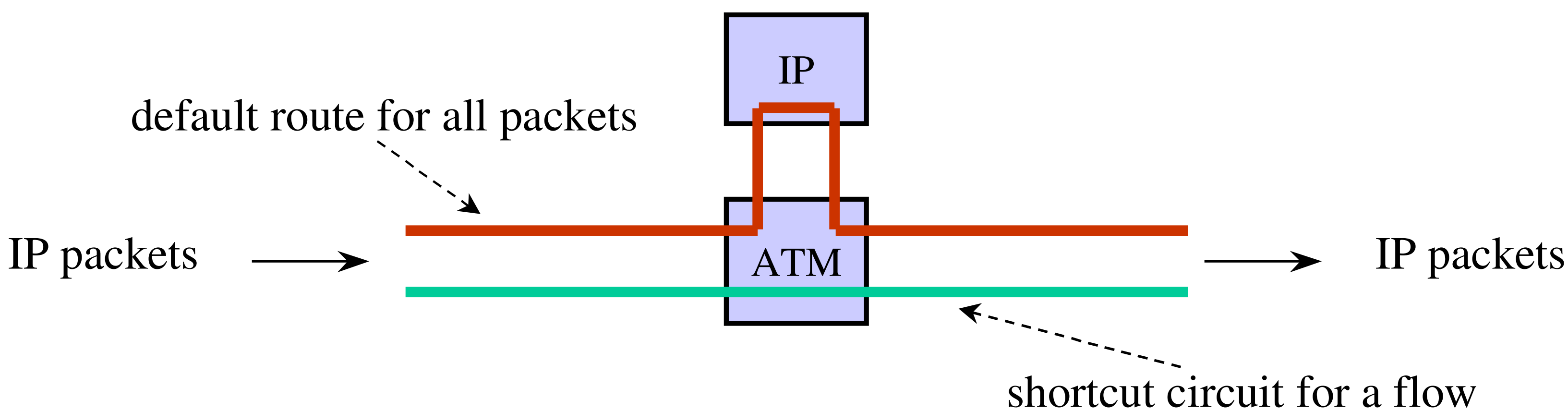
# Evaluating Network Shortcut Overheads

## Tunable architectural knobs

- Aggregation (e.g., application, host, subnet, net)
- Timeout (e.g., less than 60 seconds between packets)
- Trigger (e.g., 10 packets before shortcutting)

## Performance and cost metrics

- Percent of traffic on shortcuts (reduce software load)
- Set-up rate for shortcuts (signaling load on processor)
- Number of shortcuts (switch connection load)



# # of Simultaneous Shortcut Connections

Absolute number of shortcuts follows link speed and load

Probability density function of the number of shortcuts for

- different trigger values
- different timeout values
- different aggregation methods

Bimodal probability distribution for day and night

The # of shortcut connections is reduced by a factor of

- 3.3x by applying a 10-packet trigger
- 3.8x by aggregating at host level
- 5.8x by using both techniques
- over 90% of packets and bytes on shortcut connections

# Setup Rate for Shortcut Connections

Absolute setup rate follows link speed and load

Process: # of IP-flows is self-similar (H=0.7,0.6)

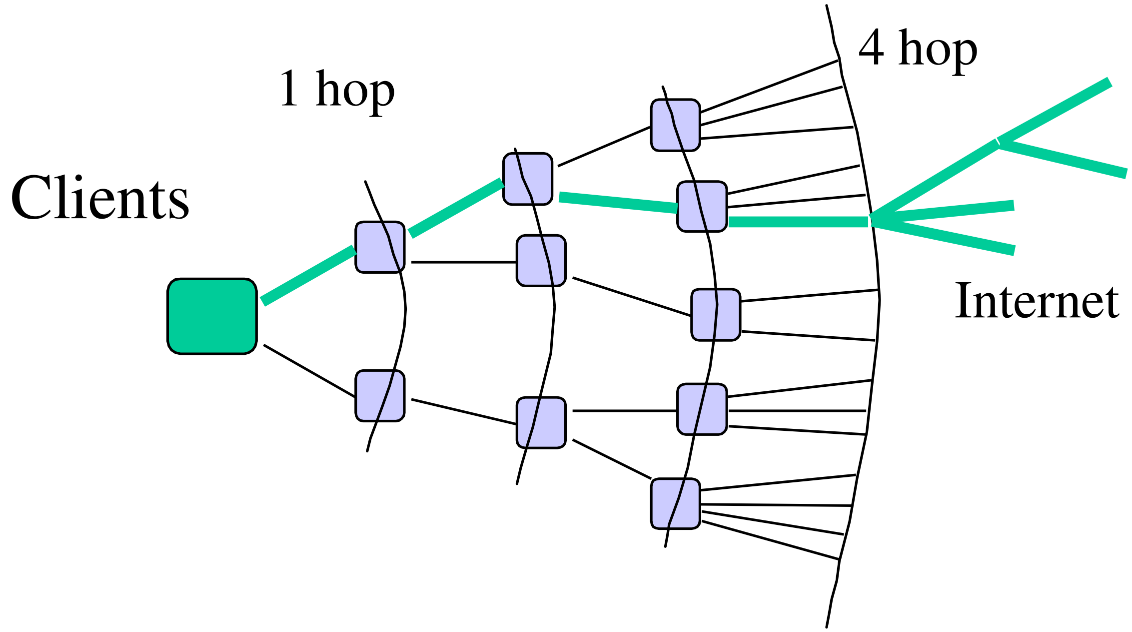
Maximum shortcut setup rate over

- different time scales (10 seconds to 3 hours)
- different trigger and timeout values
- different aggregation methods

The setup rate is reduced by a factor of

- 3.3x by applying a 10-packet trigger
- 5.2x by aggregating at host level
- 8.9x by using both techniques
- 17.8x by queueing shortcut requests up to 100 seconds

# Aggregating Traffic Along Partial Routes




Generalizing end-to-end flows to h-hop partial-routes

- Flows are usually based on the source and destination address
- But, many flows actually follow the same path through a domain
- Group related packets that travel along the same partial route
- Aggregate web responses from replicas of the same server
- Aggregate web responses from related web servers

Study: Router-to-host flows and Router-to-subnet flows

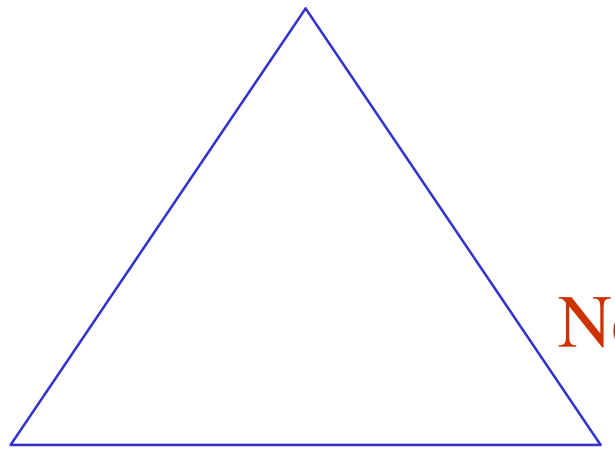
# Partial-Route Aggregation Experiments

- Evaluation methodology
  - Traceroute to over 23,000 web servers contacted in a single week
  - Exponential route fan-out (26, 71, 137, 267, 406, 916, 1508)
  - Computation of IP flows based on h-hop router address
- Router-to-host flows (single client)
  - 7-hop routes: reduce set-up by 20% and connections by 11%
  - 3-hop routes: reduce set-up by 37% and connections by 16%
- Router-to-subnet flows (group of related clients)
  - 7-hop routes: reduce set-up by 30% and connections by 15%
  - 3-hop routes: reduce set-up by 88% and connections by 65%

# Summary

## Web Protocol:

- cache misses, failed requests
- heavy-tailed file size distribution
- aggregation properties
- flow sizes == flow signature



## Trace analysis:

- long traces
- flow size distributions
  (bytes, packets, seconds)
- heavy-tailed

## Network Architecture:

MPOA, IP switching

- cost of shortcutting
- delayed setup
- route aggregation

# Conclusion

Guidelines for designing flow-switched networks for Web traffic:

- Shortcut parameters for aggregation, trigger, and timeout
- Modest aggregation and triggering without performance loss
- Reducing worst-case set-up rate through delayed shortcuts
- Reducing overheads by aggregating on partial routes